Japanese Speech Based Fuzzy Man-Machine Interface of Manipulators

Kiyotaka Izumi, Keigo Watanabe, Yuya Tamano and Kazuo Kiguchi

Department of Advanced Systems Control Engineering, Saga University, 1-Honjomachi, Saga 840-8502, Japan
(Tel: +81-952-28-8602; E-mail: izumi@me.saga-u.ac.jp)

Abstract: Recently, personal robots and home robots are developing by many companies and research groups. It is considered that a general effective interface for user of those robots is speech or voice. In this paper, Japanese speech based man-machine interface system is discussed for reflecting the fuzziness of natural language on robots, by using fuzzy reasoning. The present system consists of the derivation part of action command and the modification part of the derived command. In particular, a unique problem of Japanese is solved by applying the morphological analyzer ChaSen. The proposed system is applied for the motion control of a robot manipulator. It is proved from the experimental results that the proposed system can easily modify the same voice command to the actual different levels of the command, according to the current state of the robot.

Keywords: Speech based control, Man-machine interface, Fuzzy reasoning, Robot manipulators, Japanese analysis

1. INTRODUCTION

Research or development about personal robots and home robots is attractive in recent years. Many companies and research groups exhibit personal robots of various types such as an entertainment type, communication type, etc. at the robot dream exposition (ROBODEX 2003)[1] in Kanagawa, Japan. Personal robots will diffuse to each home widely in the near future.

The other hand, voice is used as the interface between human and robot or computer. For example, there are various amusement-type robots including a toy that can react to voice commands and some speech-recognition softwares are also sold.

Such products, as an input device with voice to the computer, are now spreading in a general home. In the future, applications of such products are further expected to be enlarged to various fields such as home-appliances, welfare instruments, etc. As for the advantages that one uses voice to an interface, it is easily point out that 1) there are few burdens to users because we need not necessarily remember the special knowledge and operations on input devices such as keyboards and mouse; 2) it is an available interface to the man who has a handicap in hand and foot, old people, and the man with whose both hands are closed by some manual operation; and 3) since it is a sociable interface ordinarily utilized by human beings, we can have an attached heart to a controlled object.

From the above backgrounds, attention has gathered also in the research of a mobile-robot control by voice commands [2]∼[5]. However, in the conventional researches of a robot control by voice, the action of the robot and the input command are the one-to-one correspondence, and consequently the future action of the robot is selected, irrespective of the present situations of the robot. Therefore, fuzziness of natural language, such that the same word produces some different interpretations according to the situation, cannot be reflected to the action of the robot. Furthermore, a trouble peculiar to Japanese language also exists in the speech recognition.

In this paper, in order to reflect the fuzziness of natural language to the action of the robot, we construct the fuzzy control system based on voice commands in Japanese, by considering the present situation of the robot. This system is composed of the derivation part of the action command and the modification part of the derived command. The former part recognizes the command by Japanese voice. The latter part derives the action command suitably modified by the fuzzy reasoning, where the fuzzy inputs consist of the voice command recognized in the former part and of the present state of the robot. The present system is applied to a redundant robot manipulator and the effectiveness is illustrated through some experiments.

2. PROBLEMS OF SPEECH BASED CONTROL

General problems regarding the robot control by speech are listed below.

Problem 1: Transformation from the voice command to the action command

The voice input can be transformed to the text by using the software of speech recognition. In Japanese, the obtained text is not separated into each word by space. This problem is different from English and other language. Therefore, first of all, the text must be divided into the unit of words in Japanese case. Also, any words have several conjugated forms. Thus, the action command has to be derived by considering the combination of each word and the conjugated form.

Problem 2: Restriction of the number of voice commands

The number of voice commands is restricted for some conventional approaches in this field. Also, users must learn such voices before using the control system. In spite of the fact that voice is a convenient input method, it is not suitable to show some inconvenience to users.

Problem 3: Limited interpretation of voice commands

The control system that the voice command and the action decision of the robot are related with one-to-one, cannot cope with real situations in which we need a flexible command depending on the state of the robot and the intention of the user. For example, consider a control system by voice

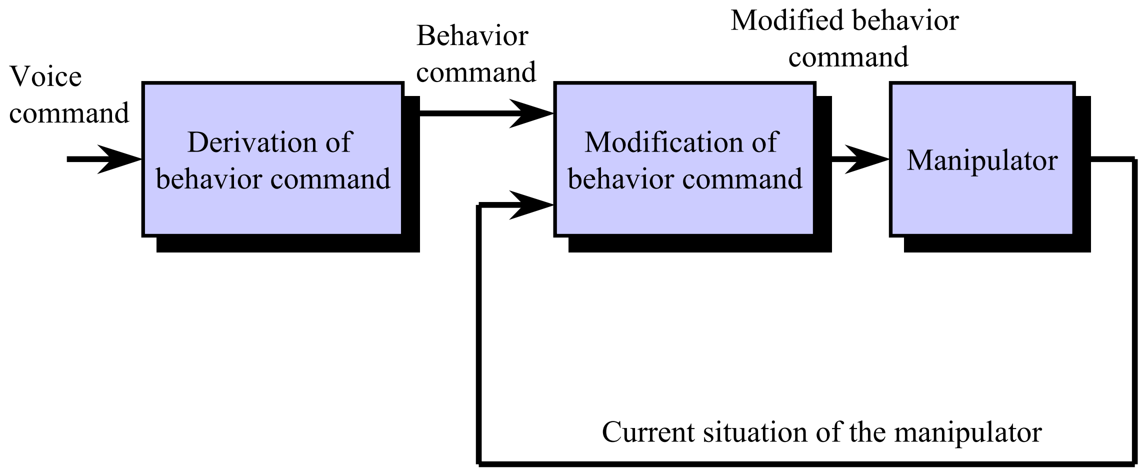

Fig. 1. Overview of the proposed system

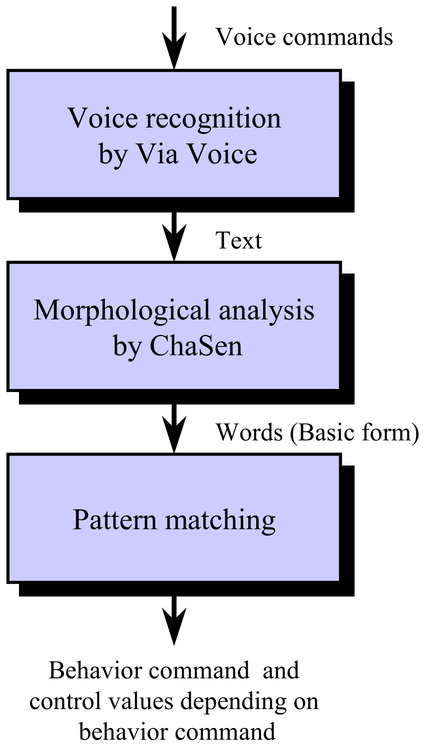

Fig. 2. Derivation part of action command

for a mobile robot. The action command is assumed to be composed of only the forward velocity of the robot. Suppose that an interpretation, "faster," was obtained by the speech recognition, and then the forward velocity was selected as 1 [m/s] by using a prespecified rule. If the current velocity is much less than 1 [m/s], then the above decision is natural for our desire. However, if the current velocity is much greater than 1 [m/s], then such a decision seems to be strange for our desire. This is attributed to the fact that the action decision was derived by a one-to-one rule, without using the present state of the robot. Thus, unless the present state of the robot is considered, a possibility that the word has wide interpretation, i.e., the fuzziness of the natural language cannot be reflected on the action decision of the robot.

3. SPEECH BASED FUZZY MAN-MACHINE INTERFACE IN JAPANESE

The propose speech based fuzzy man-machine interface in Japanese is mainly composed of the derivation part of the action command and the modification part of the derived action as shown in **Fig. 1**. The derivation part of the action command outputs the control real value for robot from the speech command obtained by microphone. Conventional speech based control can be realized using this part. In the proposed method, robot can behave reflecting the fuzziness of speech by using the modification part of behavior command.

3.1. Derivation part of the action command

The derivation part of the action command further consists of three processes, i.e., speech recognition, morpheme analysis, and pattern matching as shown in **Fig. 2**. Each process and the software used are explained below in detail.

3.1.1 Speech recognition

The voice of the user is collected through a microphone. Such voice is transformed to the text by the software of the speech recognition. The software is the ViaVoice for Windows Pro version 8 (Japanese Version) provided by IBM Co. Ltd. A Software Development Kit (SDK) to the ViaVoice is offered for the developers. In our approach, the speech recognition system is constructed by using the function of Active X in order to transform the voice to the text.

3.1.2 Morphological analysis

A Japanese morphological analysis system was introduced to solve the unique problem in Japanese, as discussed above. The system is called "ChaSen"[6], which was developed by Computational Linguistics Laboratory, Graduate School of Information Science, Nara Institute of Science and Technology (NAIST). This system divides a sentence to the words in some unit forms, where such words to be divided are determined from the parts of speech by the system. The proposed system integrates the morphological analysis system by the dynamic link library (DLL) of ChaSen.

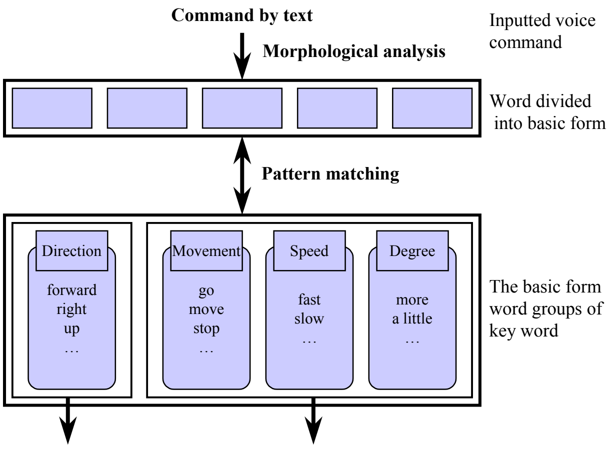

Fig. 3. Separation of command with text

Direction	Command
front	(x, y, z) = (1, 0, 0)
back	(x, y, z) = (−1, 0, 0)
left	(x, y, z) = (0, −1, 0)
right	(x, y, z) = (0, 1, 0)
up	(x, y, z) = (0, 0, 1)
down	(x, y, z) = (0, 0, −1)

↓

Sign of command

Fig. 4. Derivation of direction

		Degree				
		very	more	(none)	a little	slightly
Speed	quick	145	130	115	100	100
	fast	125	110	95	80	80
	(none)	–	–	75	–	–
	slow	25	40	55	70	70
	leisurely	5	20	35	50	50

↓

Amplitude of command

Fig. 5. Derivation of amplitude part

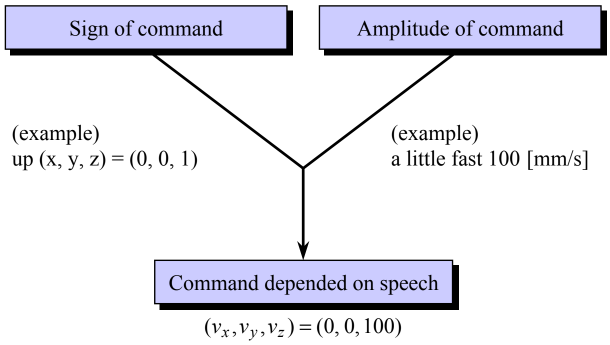

Fig. 6. Derivation using sign part and amplitude part

3.1.3 Pattern matching

The pattern matching compares the results obtained from the morphological analysis with the basic word group that represents "movement," "direction," "speed,' and "degree" as shown in **Fig. 3**. Next, the sign of the action command is obtained by the word selected in the word group of "direction" depending on the rule illustrated in **Fig. 4**. The amplitude of the action command is obtained by combination between the word selected in the word group of "speed" and the word selected in the word group of "degree" as shown in **Fig. 5**. Finally, the real value of the action command is calculated by multiplying the sign and the amplitude as shown in **Fig. 6**.

3.2. Modification part of the derived command

In the proposed method, it is considered that the fuzziness of speech shows a case in which human behaves flexibly depending on his current situation, even if the same command is given. The modification part of the derived command as shown in **Fig. 7** is discussed in this paper. This part modifies the action command derived in the derivation part so as to reflect the present state of the robot by using a fuzzy reasoning. Inputs of the fuzzy reasoning are the original robot command derived in the derivation part and the actual present state of the robot, i.e., the forward tip velocity of the robot manipulator. The fuzzy reasoning can be constructed based on humans decision making.

4. EXPERIMENTS

4.1. Robot manipulator

The proposed method is applied for the 7 degree-of-freedom redundant manipulator as shown in **Fig. 8**. The manipulator has some control mode such as position control, velocity control and direct torque control. In this experiments, the manipulator is control by velocity control mode. In other words, control inputs are the tip velocity of the manipulator. We have to set velocity command within every 200 [ms]. Unfortunately, an operator can not provide speech command; multi thread program is needed on velocity control mode. Thus, the frequency of control loop is different from that of the speech recognition loop.

4.2. Derivation part of the action command

The amplitude of action commands is set as shown in Fig. 5. Note that, the case without the word about "speed" or "degree" is considered, because actual speech command does not necessarily include such a word. However the directional command is kept, when the speech command does not include the word about "direction".

4.3. Modification part of the derived command

The fuzzy reasoning is composed of a simplified reasoning method in which the consequent part is a real number. The membership function is a type of triangle. The fuzzy rule is illustrated in **Fig. 9**. **Fig. 10** illustrates the corresponding

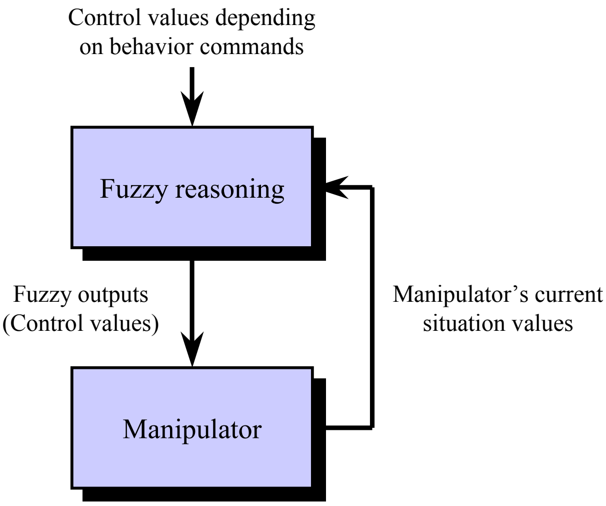

Fig. 7. Modification part of the derived command

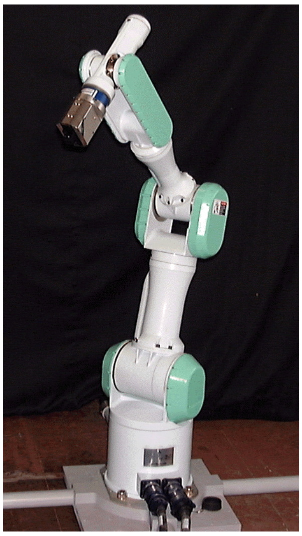

Fig. 8. Robot manipulator

		Current State				
		VS	S	M	F	VF
Voice Command	VS	VS	VS	VS	S	M
	S	VS	VS	S	M	F
	M	VS	S	M	F	VF
	F	S	M	F	VF	VF
	VF	M	F	VF	VF	VF

VS: Very Slow S: Slow M: Medium
F: Fast VF: Very Fast

Fig. 9. Fuzzy rules

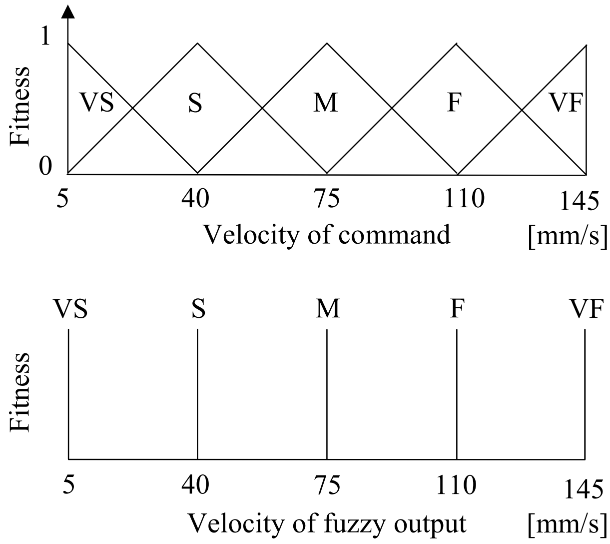

Fig. 10. Membership function

Table 1. Limitations of joint angle and angular velocity

Axis	Limitation [degree]			Maximum angular velocity [rad/s]
	Mechanical limitation	Servo limitation	Software limitation	
S1	±180	±178	±177	±1
S2	±94	±92	±91	±1
S3	±180	±175	±174	±1
E1	±143	±138	±137	±2
E2	±270	±256	±255	$\pm 2\pi$
W1	±180	±166	±155	$\pm 2\pi$
W2	none	±361	±360	$\pm 2\pi$

membership functions and the consequent parts. The rules in Fig. 9 and the design parameters in Fig. 10 are decided with mobile range of each joint as shown in **Table 1** and mobile range of work space as shown in **Fig. 11**.

4.4. Results

The speech command is given as x-axis direction of the manipulator, because we want to show the effectiveness of the proposed method clearly. The experimental result is shown in **Fig. 12**. When the same command, which is "fast" as shown in Fig. 12, is given, the amplitude of the command becomes large depending on the current velocity of the manipulator. **Table 2** illustrates the corresponding numerical data of Fig. 12. It is confirmed that the above fact also appears in Table 2. In the last row of Table 2, the speech command doesn't compose with the word of "direction" group. However, the manipulator moves to right, because the direction of the previous command is right. Finally, it is found from the result that a different action decision is obtained, even if the same speech command was given to the manipulator. Thus, observe that the present system gives a more natural action of the manipulator, like a living body who can decide the action, depending on its own internal or external state.

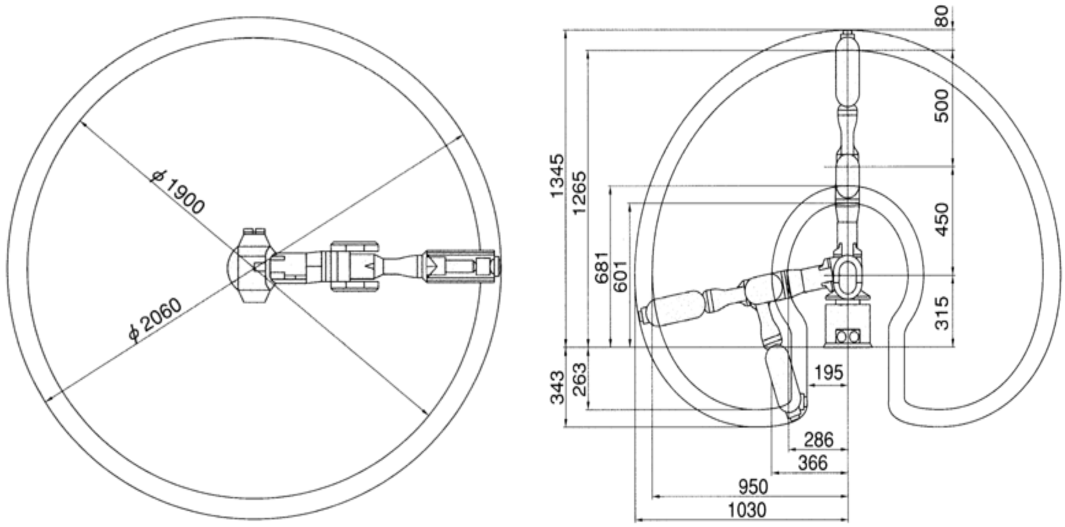

Fig. 11. Mobile range of PA10

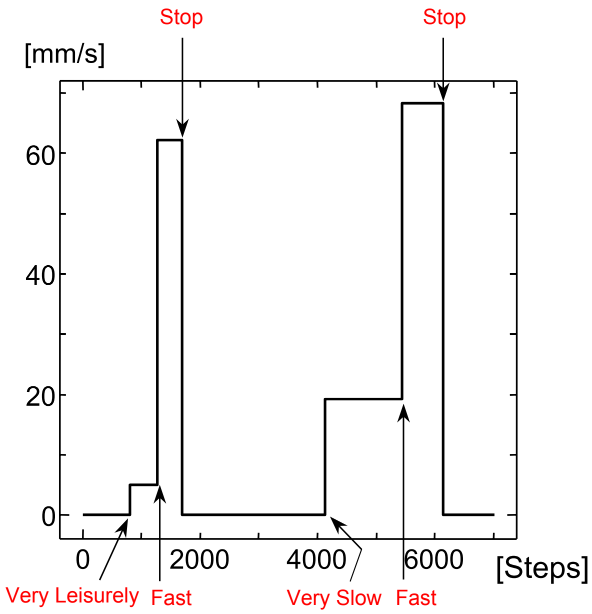

Fig. 12. Experimental result of right movement

Next, we give commands by speech in order to move up. The result is illustrated in **Fig. 13** and **Table 3**. It is found that similar result with Fig. 12 and Table 2 is obtained.

5. CONCLUSIONS

The Japanese voice control system with the fuzzy reasoning has been constructed to reflect the fuzziness of the natural language on robots. The proposed system was composed of the derivation part of the action command and the modification part of the derived command. The present system has been also applied for the motion control of the redundant robot manipulator. From the experimental results, it was found that the proposed system was able to modify the same speech command to the actual different levels of the command, according to the current state of the robot. Thus, a unique problem of Japanese was solved by applying the morphological analyzer ChaSen.

Table 2. Neumerical result of right movement

Given voice command	Velocity of command	Current velocity of PA-10	Velocity of fuzzy output
move very leisurely to right	5 mm/s	0 mm/s	5 mm/s
move quick to right	115 mm/s	5 mm/s	62 mm/s
move very slow to right	25 mm/s	0 mm/s	19 mm/s
move quick	115 mm/s	19 mm/s	68 mm/s

Table 3. Neumerical result of up movement

Given voice command	Velocity of command	Current velocity of PA-10	Velocity of fuzzy output
move up leisurely	35 mm/s	0 mm/s	35 mm/s
move quick	115 mm/s	35 mm/s	75 mm/s
move very quick	145 mm/s	75 mm/s	105 mm/s

References

[1] URL: http://www.robodex.org

[2] K. Komiya, K. Moita, K. Kakegawa, and K. Kurosu, "Guidance of a wheelchair by voice," in *Proc. of IECON'2000*, 2000, pp. 102–107.

[3] K. Pulasinghe, K. Watanabe, K. Kiguchi, and K. Izumi, "A novel modular fuzzy neural controller driven by natural language commands," in *Proc. of the 40th SICE*

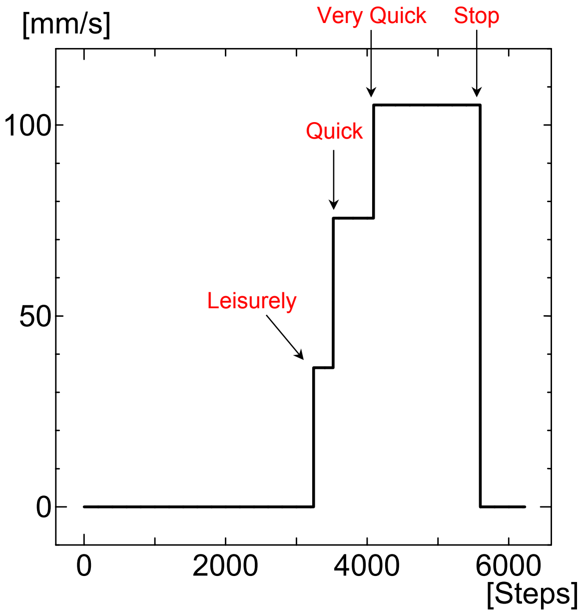

Fig. 13. Experimental result of up movement

Annual Conference (SICE2001), Int. Session Papers, 2001, 312C-4.

[4] K. Pulasinghe, K. Watanabe, K. Kiguchi and K. Izumi, "Modular fuzzy neural controller driven by voice commands," in *Proc. of ICCAS2001*, 2001, pp. 194–197.

[5] C. T. Lin and M. C. Kan, "Adaptive fuzzy command acquisition with reinforcement learning," *IEEE Trans. on Fuzzy Systems*, vol. 6, no. 1, pp. 102–121, 1998.

[6] Y. Matsumoto, A. Kitauchi, T. Yamashita, Y. Hirano, H. Matsuda, K. Takaoka, and M. Asahara, "Morphological analysis system ChaSen version 2.2.9 manual," http://chasen.aist-nara.ac.jp/chasen/doc/chasen-2.2.9.pdf, 2002.